Half-Simple Symmetric Venn Diagrams

Charles E. Killian
Department of Computer Science,
Duke University, Durham, NC
ckillian@cs.duke.edu

Frank Ruskey *
Department of Computer Science,
University of Victoria, Victoria, BC

Carla D. Savage †
Department of Computer Science,
North Carolina State University, Raleigh, NC
savage@csc.ncsu.edu

Mark Weston
Department of Computer Science,
University of Victoria, Victoria, BC
mweston@cs.uvic.ca

Submitted: Sep 9, 2004; Accepted: Oct 13, 2004; Published: Jan 1, 2005
Mathematics Subject Classifications: 05A10, 05A16, 06A07, 06E10

Abstract

A Venn diagram is simple if at most two curves intersect at any given point. A recent paper of Griggs, Killian, and Savage [Elec. J. Combinatorics, Research Paper 2, 2004] shows how to construct rotationally symmetric Venn diagrams for any prime number of curves. However, the resulting diagrams contain only $\binom{n}{\lfloor n/2 \rfloor}$ intersection points, whereas a simple Venn diagram contains $2^n - 2$ intersection points. We show how to modify their construction to give symmetric Venn diagrams with asymptotically at least 2^{n-1} intersection points, whence the name "half-simple."

1 Introduction

Following Grünbaum [5], a *Venn diagram* for n sets is a collection of n simple closed curves in the plane, $\{\Theta_1, \Theta_2, \ldots, \Theta_n\}$, with the property that for each $S \subseteq \{1, 2, \ldots, n\}$ the region

$$\bigcap_{i \in S} \mathrm{int}(\Theta_i) \quad \cap \quad \bigcap_{i \notin S} \mathrm{ext}(\Theta_i)$$

is nonempty and connected. Here $\mathrm{int}(\Theta_i)$ and $\mathrm{ext}(\Theta_i)$ denote the open interior and open exterior, respectively, of Θ_i. A Venn diagram is *simple* if no 3 curves have a common point of intersection. In a Venn diagram the curves are assumed to be finitely intersecting. A Venn diagram is *rotationally symmetric* if there is a point p such that each of the rotations

*Research supported in part by NSERC grant 0-GP-000337999
†Research supported in part by NSF grant DMS-0300034 and NSA grant MDA 904-01-0-0083

of Θ_1 about p by an angle of $2\pi i/n$, $0 \le i \le n-1$, coincides with one of the curves $\Theta_1, \Theta_2, \ldots, \Theta_n$. A Venn diagram is *monotone* if every region enclosing k curves, with $0 < k < n$, is adjacent to a region enclosing $k-1$ curves and a region enclosing $k+1$ curves. Monotone Venn diagrams are precisely those that can be drawn with all curves convex, as shown in [1].

Venn diagrams for n sets with rotational symmetry cannot exist unless n is prime (Henderson [8]) and it is shown in [4] that they do exist for all prime n, by a general construction. The symmetric diagrams in [4] contain $\binom{n}{\lfloor n/2 \rfloor}$ intersection points, with exactly n points of intersection through which all curves pass. Such diagrams were introduced by Ruskey and Chow [9], who provided examples of them for $n = 5$ and $n = 7$. In [7] Hamburger constructed a symmetric Venn diagram for $n = 11$ with this property.

It follows from Euler's formula ($V - E + F = 2$) that a simple Venn diagram has $2^n - 2$ vertices. The diagrams constructed in [4], both symmetric and non-symmetric, are monotone, and among monotone Venn diagrams they contain the least number of vertices, namely:

$$\binom{n}{\lfloor \frac{n}{2} \rfloor} \sim \frac{2^n}{\sqrt{n}}. \tag{1}$$

This was shown in [2] to be minimum for monotone Venn diagrams.

Simple monotone non-symmetric diagrams exist for all n, but simple symmetric Venn diagrams are known only for $n = 3, 5$ and 7: see [9] for some examples. So, we are now motivated to ask if we can we find *simple* symmetric Venn diagrams for all prime n, or at least ones that are "more simple", where as a measure of simplicity, we use the number of vertices in the Venn diagram.

In this paper we show that for n prime, we can in fact add vertices to the diagrams in [4] to produce symmetric diagrams in which the number of vertices is asymptotically at least 2^{n-1}. The technique used in [4] makes use of some novel observations about the Greene-Kleitman symmetric chain decomposition of the Boolean lattice [3]. The paper [4] also includes a construction, for any n, of monotone (non-symmetric) Venn diagrams with the minimum number of vertices. We show that when the same method of "adding vertices" is applied to this case, the resulting Venn diagrams have, surprisingly, *exactly* 2^{n-1} vertices. In both cases, to accomplish this we manipulate the dual graph of the diagram, which has 2^n vertices, each of which is identified with a bitstring of length n. Since vertices in a graph correspond to faces in the dual, we need to modify the construction so that the dual has more faces.

The survey [9] introduced the idea of a *separable vertex* as a way of simplifying a Venn diagram by introducing more vertices without destroying the Venn diagram property. Adding an edge to this dual graph corresponds to separating a vertex of the Venn diagram into two vertices by pulling some curves out of the vertex. A face in the dual that cannot be subdivided into several smaller faces corresponds to a vertex that is not separable in the Venn diagram; if all vertices are not separable the diagram is termed *rigid*. Simple diagrams are trivially rigid. Thus the technique of adding as many faces as possible to the dual graph corresponds to separating as many vertices as possible in the Venn diagram until it becomes rigid.

The next section describes the Venn diagram construction from [4] which we refer to as *the GKS construction*, and uses an example to illustrate the general idea. In Section 3 we prove a theorem used to subdivide faces in dual graphs into many 4-faces, each of which corresponds to the intersection of two curves in the Venn diagram. In the remaining sections, this theorem is then applied to the non-symmetric monotone diagrams and then the symmetric diagrams to calculate how many new faces, and thus vertices, can be added in each case to the diagrams.

2 The GKS Construction

The Hasse diagram of the Boolean lattice $\mathcal{B}_n$, when viewed as a graph, is isomorphic to the n-cube, the graph whose vertices are the n-bit strings, with two vertices joined by an edge when they differ in only one bit. The isomorphism maps a set $S \subseteq \{1, \ldots, n\}$ to the n-bit string with 1 in position i iff $i \in S$. The *weight* of a bit string is the number of ones it contains.

A subgraph of the n-cube is called *monotone* if every vertex of weight d is adjacent to a vertex of weight $d+1$ (if $d < n$) and a vertex of weight $d-1$ (if $d > 1$).

The GKS construction is based on the following Theorem, proved in [4].

Theorem 1 *If G is a plane, monotone, spanning subgraph of the n-cube, the dual of G is a (monotone) Venn diagram.*

To construct "simpler" Venn diagrams, the plan to is to get more *vertices* in the Venn diagram by adding more *edges* in the dual. We start with an intermediate construction that works for *all* n to make the GKS construction "simpler" (but not symmetric) and then show how to modify this when n is prime to make it symmetric as well. In the remainder of this section, we review the GKS construction.

In a binary string, regard each '1' as a right parenthesis and each '0' as a left parenthesis and match parentheses in the usual way. For example, in the string

$$1\ 0\ 0\ 1\ 0\ 0\ 1\ 1\ 1\ 1\ 0\ 0\ 1\ 0$$

the '1' bits in positions 4, 7, 8, 9, and 13 are matched, respectively, with the '0' bits in positions 3, 5, 6, 2, 12. The '1' in position 10 is unmatched and the '0' bits in positions 11 and 14 are unmatched.

For every $n > 0$, define the rooted tree T_n to be the tree whose nodes are the n-bit strings with no unmatched 1, and where the parent of node x, $p(x)$, is obtained from x by changing the last 1 in x to 0. Note $p(x)$ is in T_n when x is, since if x has no unmatched 1, the same must be true of $p(x)$. See Figure 1 for T_5.

Given T_n, we now grow each node x of T_n into a *chain* C_x of n-bit strings using the *Greene-Kleitman successor rule* [3]:

> Starting with a string x with no unmatched 1, change the first unmatched 0 to 1 to get its successor, y. Change the first unmatched 0 in y to 1 (if any) to get its successor. Continue until a string with no unmatched 0 is reached.

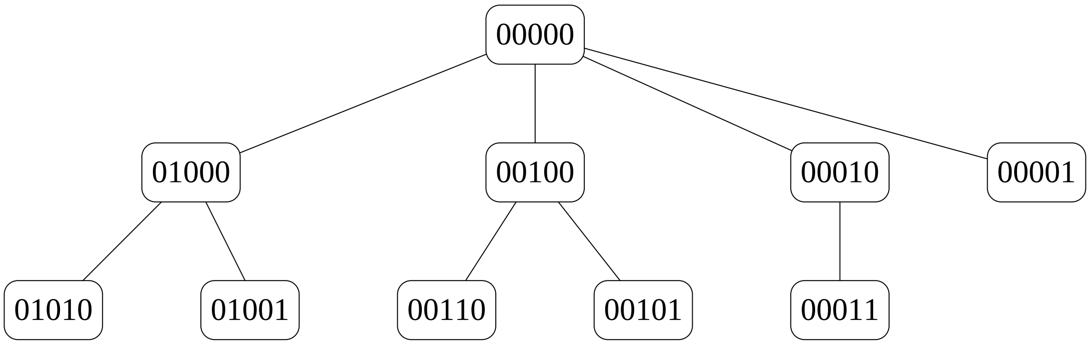


Figure 1: The rooted tree T_5 of 5-bit strings with no unmatched 1.

As an example, by this rule node 00100 in T_5 is expanded into the chain C_{00100}:

$$C_{00100}: \quad 00100 \to 10100 \to 10110 \to 10111$$

and the complete list of chains for nodes in T_5 is:

$$\begin{array}{ll}
C_{00000}: & 00000 \to 10000 \to 11000 \to 11100 \to 11110 \to 11111 \\
C_{01000}: & 01000 \to 01100 \to 01110 \to 01111 \\
C_{01010}: & 01010 \to 01011 \\
C_{01001}: & 01001 \to 01101 \\
C_{00100}: & 00100 \to 10100 \to 10110 \to 10111 \\
C_{00110}: & 00110 \to 00111 \\
C_{00101}: & 00101 \to 10101 \\
C_{00010}: & 00010 \to 10010 \to 11010 \to 11011 \\
C_{00011}: & 00011 \to 10011 \\
C_{00001}: & 00001 \to 10001 \to 11001 \to 11101
\end{array}$$

Greene and Kleitman showed in [3] that this gives a symmetric chain decomposition of the Boolean lattice. That is, (1) each of the resulting paths in the n-cube is a *chain* in $\mathcal{B}_n$: each element is covered in $\mathcal{B}_n$ by its successor; (2) in each chain, the weights of the first and last elements sum to n; and (3) every n-bit string is in exactly one of the chains.

By definition of T_n, the first elements of C_x and $C_{p(x)}$ differ in only one bit. In [4] it is shown that the last elements of C_x and $C_{p(x)}$ also differ in only one bit and the chains are used to get a Venn diagram as follows.

We use the chain decomposition derived from T_n to build a plane graph $P(T_n)$ that forms the dual of the final Venn diagram. Embed the chains $\{C_x | x \in T_n\}$ vertically in the plane, one unit apart and centered about some horizontal line, by *preorder* in T_n; that is, (1) for every x, C_x precedes C_y if y is a descendant of x in T_n and (2) regarding the children of x as ordered $x_1, ..., x_t$ for $1 \leq i \leq t-1$, chains for descendants of x_i must appear before chains for descendants of x_{i+1}. (See Figure 2, ignoring the green edges). In addition, for each x of positive weight, whenever x is the first child of its parent $p(x)$ we include the edge joining the first elements of C_x and $C_{p(x)}$ and the last elements of C_x

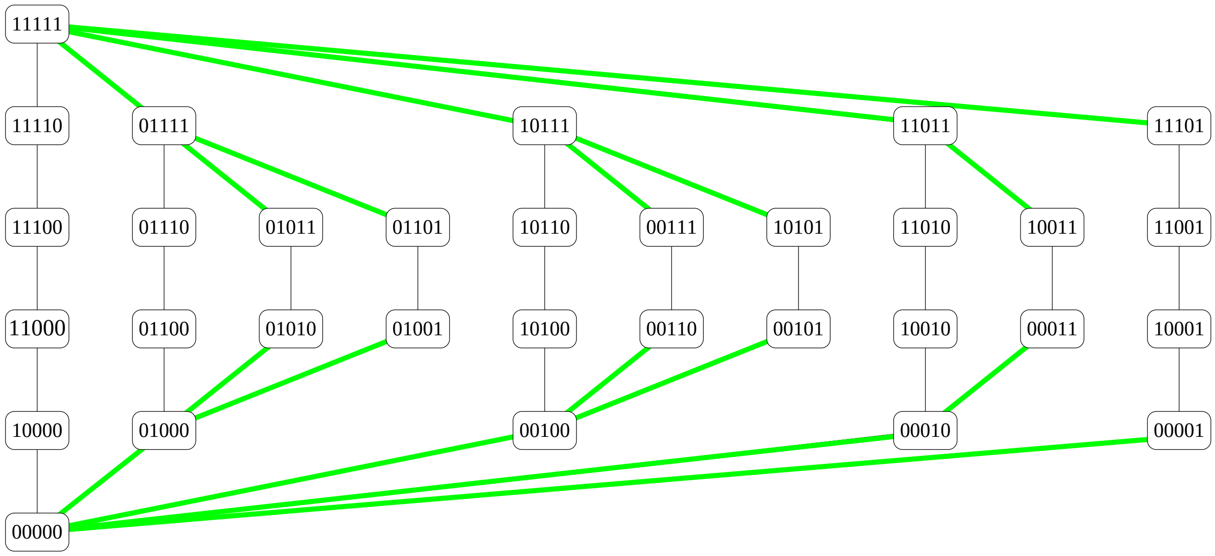


Figure 2: The plane graph $P(T_5)$, with attaching edges drawn in green.

and $C_{p(x)}$ (these edges must always exist, as noted above). Henceforth we refer to these edges, shown as the green edges in Figure 2, as the *attaching edges* for C_x.

As shown in [4], the resulting graph $P(T_n)$ is a plane, monotone, spanning subgraph of the n-cube, so by Theorem 1, its dual is a Venn diagram. Figure 3 illustrates the process of taking the dual of $P(T_4)$, and Figure 4 shows the resulting Venn diagram for 4 sets. In this construction, the number of vertices in the resulting Venn diagram will be the same as the number of chains in the symmetric chain decomposition of the Boolean lattice $\mathcal{B}_n$, which is $\binom{n}{\lfloor n/2 \rfloor}$.

Moving now to prime n, the notion of block code for strings under rotation was introduced in [4] and was the key breakthrough in showing the existence of symmetric Venn diagrams for all prime n. Define the *block code* $\beta(x)$ of a binary string x as follows. If x starts with 0 or ends with 1, then $\beta(x) = (\infty)$. Otherwise, x can be written in the form:

$$x = 1^{a_1}0^{b_1}1^{a_2}0^{b_2}\cdots 1^{a_t}0^{b_t}$$

for some $t > 0$, where $a_i > 0$, $b_i > 0$, $1 \le i \le t$, in which case,

$$\beta(x) = (a_1 + b_1, a_2 + b_2, \ldots, a_t + b_t).$$

The n *rotations* of x, where x has length n, are the n strings reachable by applying the circular permutation $(12\cdots n)$ to x. As an example, the block codes of the string 1110101100010 and all of its rotations are shown below.

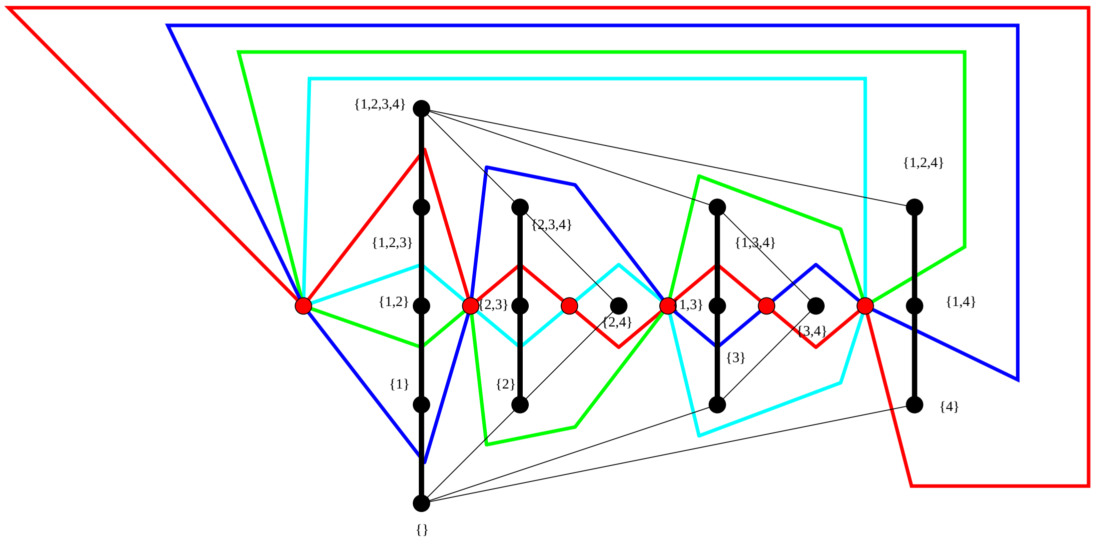


Figure 3: Construction of dual of $P(T_4)$.

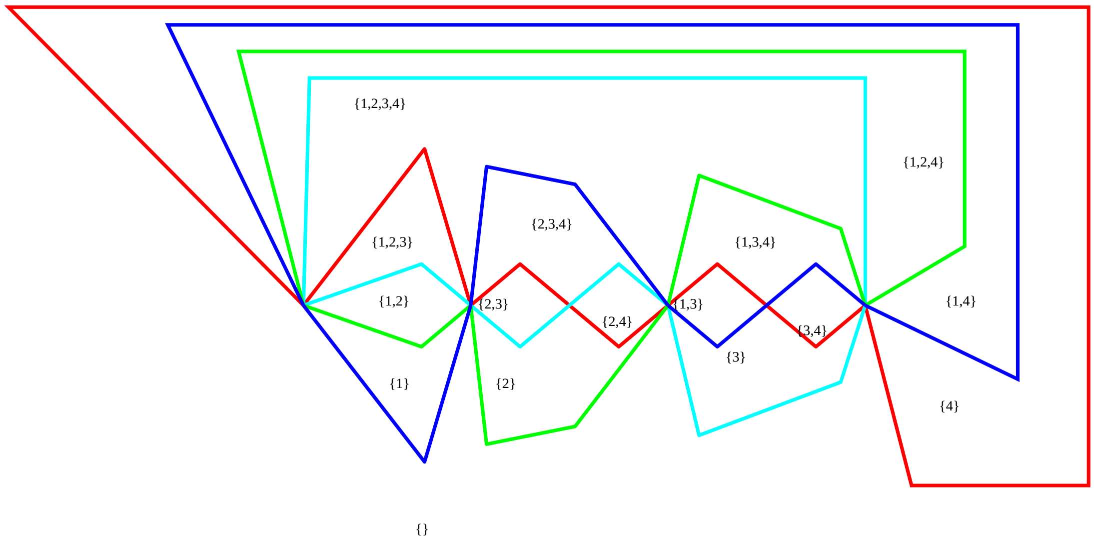


Figure 4: Venn diagram for $n = 4$.

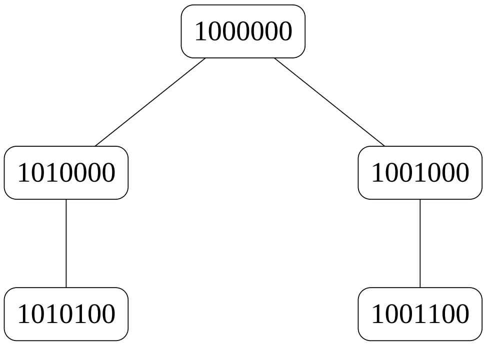


Figure 5: The rooted tree S_7 of 7-bit rotational equivalence class representatives with one unmatched 1.

bit string	block code	bit string	block code
1110101100010	$(4, 2, 5, 2)$	1100010111010	$(5, 2, 4, 2)$
0111010110001	(∞)	0110001011101	(∞)
1011101011000	$(2, 4, 2, 5)$	1011000101110	$(2, 5, 2, 4)$
0101110101100	(∞)	0101100010111	(∞)
0010111010110	(∞)	1010110001011	(∞)
0001011101011	(∞)	1101011000101	(∞)
1000101110101	(∞)		

When n is prime, every n-bit string, other than $00\cdots00$ and $11\cdots11$, has n distinct rotations. Similarly, it is shown in [4] that when n is prime, no two different rotations of an n-bit string can have the same *finite* block code. Assuming that block codes are ordered lexicographically, in each equivalence class of n-bit strings under rotation, the unique string with minimum block code can be chosen as the representative.

For prime n, define the rooted tree S_n to be the tree whose nodes are the n-bit strings x with exactly *one* unmatched 1 and with $\beta(x) \leq \beta(y)$ for any rotation y of x. Note that the unmatched 1 must appear in the leftmost position of x. The parent of node x, $p(x)$, is obtained from x by changing the last 1 in x to 0. We note that $p(x)$ is in S_n when x is, since it is shown in [4] that (i) if x has exactly one unmatched 1, the same is true for $p(x)$ and (ii) if $\beta(x)$ is minimal under all rotations of x, then $\beta(p(x))$ is also. See Figure 5 for S_7.

Given S_n, we now grow each node x of S_n into a *chain* C_x of n-bit strings using the following variation of the Greene-Kleitman construction[3]:

> Start with a string x which has one unmatched 1 and lexicographically smallest block code among all of its rotations (*i.e.*, a node in S_n). If there is more than one unmatched 0 in x, change the first unmatched 0 to 1 to get its successor, y. If there is more than one unmatched 0 in y, change the first unmatched 0 in y to 1 to get its successor. Continue until a string with only one unmatched 0 is reached.

Note that a node x and its successor y have the same block code, so if x has the minimum block code among all of its rotations, then so does y. Thus every element of C_x is the (minimum block code) representative of its equivalence class under rotation.

The list of chains for nodes in S_7 is:

$$\begin{array}{ll}
C_{1000000}: & 1000000 \to 1100000 \to 1110000 \to 1111000 \to 1111100 \to 1111110 \\
C_{1010000}: & 1010000 \to 1011000 \to 1011100 \to 1011110 \\
C_{1010100}: & 1010100 \to 1010110 \\
C_{1001000}: & 1001000 \to 1101000 \to 1101100 \to 1011110 \\
C_{1001100}: & 1001100 \to 1101100 \to 1101110.
\end{array}$$

It is shown in [4] that this gives a symmetric chain decomposition of the subposet of the Boolean lattice induced by the representatives of equivalence classes of n-bit strings under rotation (with finite block code). That is, (1) each of the resulting paths is a *chain* in $\mathcal{B}_n$, (2) each element of each chain is the (minimum block code) representative of its equivalence class under rotation, (3) in each chain, the weights of the first and last elements sum to n; and (4) for every n-bit string $x \in \mathcal{B}_n - \{0^n, 1^n\}$, the rotation of x with lexicographically block code smallest rotation is in exactly one of the chains.

By definition of S_n, the first elements of C_x and $C_{p(x)}$ differ in only one bit. In [4] it is shown that the last elements of C_x and $C_{p(x)}$ also differ in only one bit and the chains are used to get a Venn diagram as follows. As before, embed the chains $\{C_x | x \in S_n\}$ vertically in the plane, one unit apart and centered about some horizontal line, by *preorder* in S_n, including the attaching edges for each C_x. As shown in [4], the resulting graph, which we call $P(S_n)$, is a plane, monotone, subgraph of the n-cube, but it only contains about $1/n$ of the the vertices of the n-cube. What we need is a plane, monotone, *spanning* subgraph of the n-cube, so that by Theorem 1, its dual is a Venn diagram. In addition, we want rotational symmetry.

This is done in [4] by constructing a graph $R(P(S_n))$ as follows. Start by making n copies of $P(S_n)$, and adding n copies each of the vertices 0^n and 1^n adjacent to each of the copies of 10^{n-1} and $1^{n-1}0$, which start and finish the longest chain. In the ith copy of $P(S_n)$, each vertex x is replaced by the rotated vertex $\sigma^i(x)$, where $\sigma^i(x_1 x_2 \cdots x_n) = x_{i+1} x_{i+2} \cdots x_n x_1 \cdots x_i$. Each copy is embedded (symmetrically) in a $1/n$-th pie slice in the plane, with the vertices 1^n coinciding at the center and the vertices 0^n coinciding at the point at infinity. Now, $R(P(S_n))$ is a plane, monotone, *spanning* subgraph of the n-cube, so by Theorem 1, its dual is a Venn diagram. Finally, the dual of $R(P(S_n))$ is constructed, preserving the symmetry. The process is illustrated in Figure 6 for $n = 5$, proceeding from $P(S_5)$ in Figure 6a through $R(P(S_5))$ in 6b to the dual of $R(P(S_5))$ in 6d.

This construction yields rotationally symmetric Venn diagrams in which the number of vertices is the same as the number of chains in a symmetric chain decomposition of $\mathcal{B}_n$, $\binom{n}{\lfloor n/2 \rfloor}$. In the next section we show a systematic way to add faces to the dual, and thereby vertices to the Venn diagram.

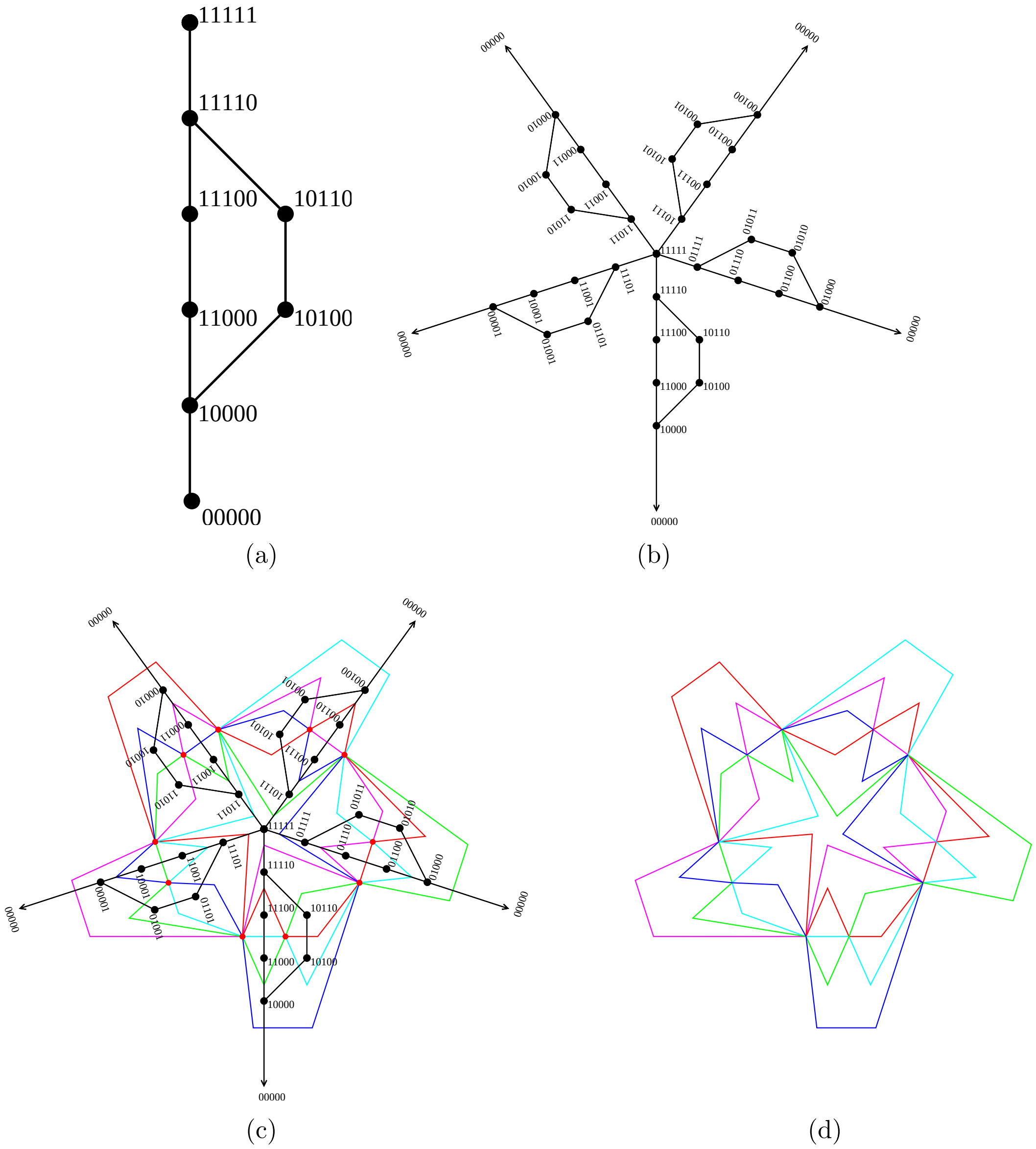


Figure 6: Building the symmetric 5-Venn diagram from $P(S_5)$.

3 Quadrangulating Faces

In this section we show that in the chain graph, $P(T_n)$, any face formed by a chain and its first child can be *quadrangulated* (*i.e.* decomposed into faces bordered by 4 edges, that is, *4-faces*) by adding non-crossing edges of the n-cube. Quadrangulating all such faces results in a plane monotone spanning subgraph $Q(T_n)$ of the n-cube. By Theorem 1, the dual of $Q(T_n)$, $Q(T_n)^*$, is a Venn diagram. Since $Q(T_n)$ has more faces than $P(T_n)$, the Venn diagram $Q(T_n)^*$ has more vertices than $P(T_n)^*$. Similarly, for prime n we show that in the chain graph $P(S_n)$, any face formed by a chain and its first child can be quadrangulated by adding non-crossing edges of the n-cube and use this to get a symmetric Venn diagram with more vertices. In Sections 4 and 5 we count how many vertices are added to the Venn diagrams which result from quadrangulating these faces in $P(T_n)$ and $P(S_n)$.

Let $|C|$ denote the length of a chain C, that is, its number of edges.

Theorem 2 *Let $w \neq 0^n$ be a node in T_n and let x be its parent. If chains C_x, C_w are embedded consecutively in $P(T_n)$, the face bounded by the chains C_w, C_x, and the attaching edges of C_x can be quadrangulated into $|C_w|+1$ 4-faces by adding $|C_w|$ edges of the n-cube (as shown in Figure 7).*

Proof. Since w is a node of T_n, w has no unmatched 1. Let b be the position of the last 1 in w and let a be the position of the 0 to which it is matched. Then w has the form $w = y10^{n-b}$ and $x = y0^{n-b+1}$. Note (i) that $a < b$ and w has no unmatched 0 between a and b (else the 1 in position b would have preferred it to the 0 in position a.) Also note (ii) that in x position a and b both contain unmatched 0 bits (there is no 1 to the right of b in x; in w, no 1 in y matched to the 0 in position a, so this remains true in x.) Finally, note (iii) that if $U_0(y)$ denotes the set of positions of the unmatched 0 bits in a string y, then $U_0(x)$ is the disjoint union $U_0(x) = U_0(w) \cup \{a, b\}$.

Let $U_0(w) = \{u_1, u_2, \ldots\ldots u_m\}$, where $u_1 < u_2 < \cdots < u_m$. By (i) above, there exists i, $0 \leq i \leq m$ such that $u_1 < u_2 < \cdots < u_i < a < b < u_{i+1} < \cdots < u_m$. By (iii), $U_0(x) = \{u_1, u_2, \ldots, u_i, a, b, u_{i+1}, u_{i+2}, \ldots, u_m\}$.

For $S \subseteq \{1, \ldots, n\}$, Define I_S to be the n-bit string with i-th bit '1' iff $i \in S$. Then the chain grown from w by the Greene-Kleitman successor rule (change first unmatched 0 to 1) is the chain of length m:

$$C_w : \quad C_w(0), C_w(1), \ldots, C_w(m),$$

where $C_w(0) = w$ and for $1 \leq j \leq m$,

$$C_w(j) = w + I_{\{u_1, u_2, \ldots, u_j\}}.$$

Here '+' denotes bitwise or. The chain grown from x is:

$$C_x : \quad C_x(0), C_x(1), \ldots, C_x(m+2),$$

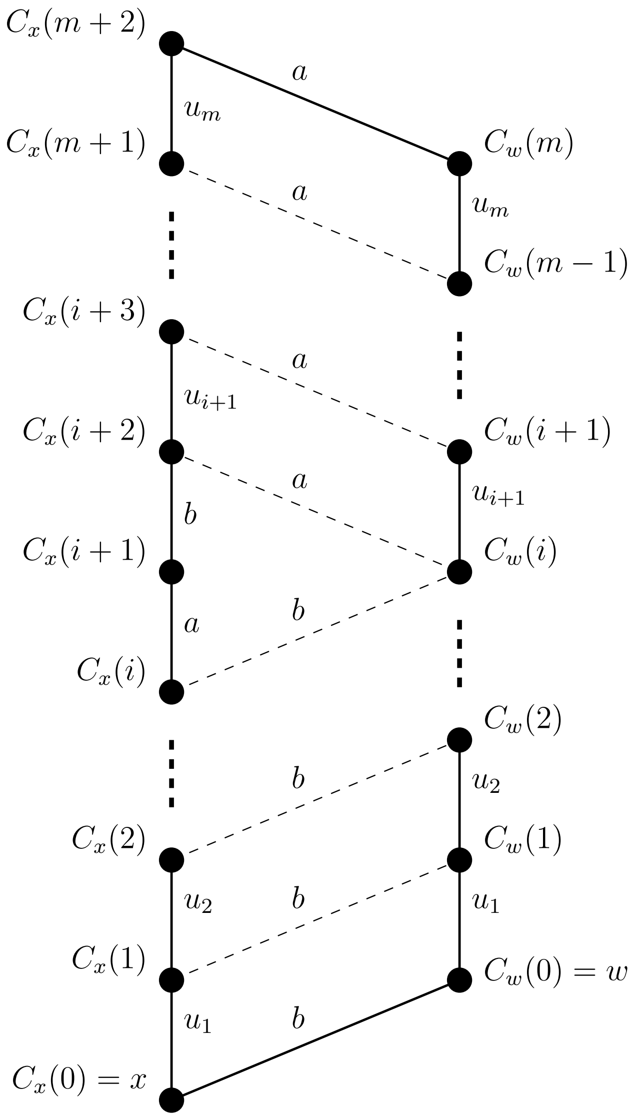


Figure 7: Quadrangulation of the face bordered by the chain C_w, its parent chain C_x, and the two attaching edges.

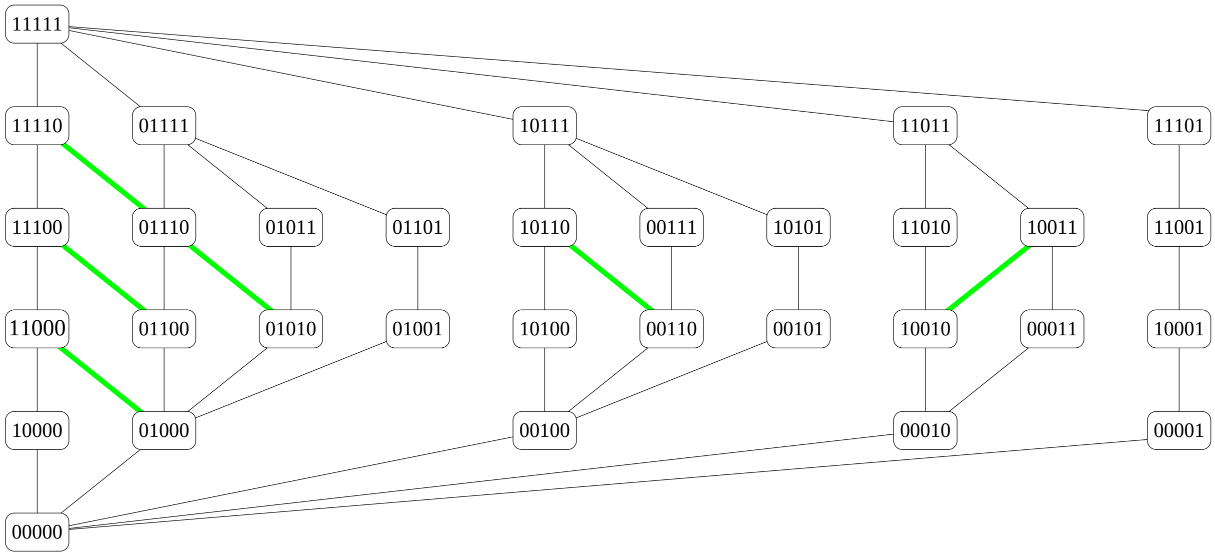


Figure 8: The plane graph $P(T_5)$ with quadrangulated faces.

where

$$C_x(j) = \begin{cases} x & \text{if } j = 0 \\ x + I_{\{u_1,u_2,\ldots,u_j\}} & \text{if } 1 \leq j \leq i \\ x + I_{\{u_1,u_2,\ldots,u_{j-1}\}} + I_{\{a\}} & \text{if } j = i+1 \\ x + I_{\{u_1,u_2,\ldots,u_{j-2}\}} + I_{\{a,b\}} & \text{if } i+2 \leq j \leq m+2 \end{cases}$$

Then for $0 \leq j \leq i$, $C_w(j)$ and $C_x(j)$ are adjacent in the n-cube, since they differ only in bit b, where w and x differ. For $i \leq j \leq m$, $C_w(j)$ and $C_x(j+2)$ are adjacent in the n-cube: they differ only in bit a, since

$$C_x(j+2) = x + I_{\{u_1,u_2,\ldots,u_j\}} + I_{\{a,b\}} = w + I_{\{u_1,u_2,\ldots,u_j\}} + I_{\{a\}} = C_w(j) + I_{\{a\}}.$$

Finally, if chains C_x, C_w are embedded consecutively in $P(T_n)$, the m edges

$$(C_x(1), C_w(1)),\ (C_x(2), C_w(2)),\ \ldots\ ,\ (C_x(i), C_w(i)),\ (C_x(i+2), C_w(i)),$$
$$(C_x(i+3), C_w(i+1)),\ \ldots\ ,\ (C_x(m+1), C_w(m-1))$$

can be added without crossings to create $m+1$ faces interior to the original face. $\square$

Figure 8 shows the plane graph that results if all faces of $P(T_5)$ bounded by chains corresponding to a node and its first child in T_5 are quadrangulated; the added edges are shaded in the figure.

We get a similar result for the plane graph $P(S_n)$.

Theorem 3 *Let $w \neq 10^{n-1}$ be a node in S_n and let x be its parent. If chains C_x, C_w are embedded consecutively in $P(S_n)$, the face bounded by the chains C_w, C_x, and the attaching edges of C_x can be quadrangulated into $|C_w|+1$ 4-faces by adding $|C_w|$ edges of the n-cube.*

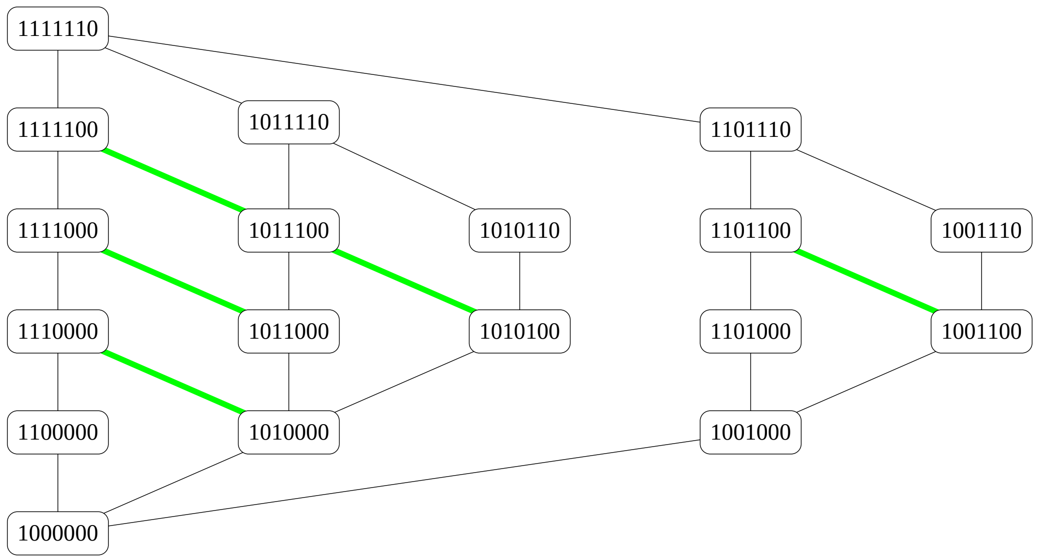


Figure 9: The plane graph $P(S_7)$ with quadrangulated faces.

Proof. The same proof works as for Theorem 2 after observing that for Theorem 3, if $U_0(w) = \{u_1, u_2, \ldots u_m\}$, then C_w is the chain:

$$C_w : \quad C_w(0), C_w(1), \ldots, C_w(m-1).$$

□

Figure 9 shows the plane graph that results if all faces of $P(S_7)$ bounded by chains corresponding to a node and its first child in S_7 are quadrangulated; the added edges are shaded in the figure.

4 Half-Simple Venn Diagrams for All n

In this section we show that in $P(T_n)$, quadrangulating every face corresponding to a node and its first child gives a Venn diagram with 2^{n-1} vertices. Since quadrangulating preserves the property that $P(T_n)$ is a plane, monotone, spanning subgraph of the n-cube, the dual is still a Venn diagram. It remains only to count the number of faces added to $P(T_n)$ by quadrangulating.

We make the following definitions.

- Let N_n be the total number of nodes in T_n. Each node corresponds to one chain in the symmetric chain decomposition of $\mathcal{B}_n$, and the number of chains is just the number of elements at the middle level of $\mathcal{B}_n$: $N_n = \binom{n}{\lfloor \frac{n}{2} \rfloor}$.

- Let $N_n(d)$ denote the number of nodes of weight d in T_n. (Nodes of weight d are at depth d in T_n.) Then $N_n(d) = 0$ if $d \geq \lfloor \frac{n}{2} \rfloor$. Since the chain starters at level d

in the Boolean lattice are those strings of weight d which do not belong to chains started by strings of smaller weight, we have

$$N_n(d) = \begin{cases} \binom{n}{d} - \binom{n}{d-1} & 0 < d \leq \lfloor \frac{n}{2} \rfloor \ , \\ 1 & \text{if } d = 0, \\ 0 & \text{otherwise.} \end{cases} \tag{2}$$

Note that $\binom{n}{d} - \binom{n}{d-1} = \frac{n-2d+1}{n-d+1}\binom{n}{d}$ is a member of the well-studied set of "ballot numbers".

- Let $\overline{L}_n(d)$ denote the number of nonleaves of weight d in T_n.

If a node x has a child in T_n, and if w is its first child, then in the plane graph $P(T_n)$, C_w is embedded immediately to the right of C_x. By Theorem 2, we can add $|C_w| = |C_x| - 2$ edges to $P(T_n)$ to quadrangulate the face bounded by C_x, C_w, and the attaching edges of C_w. Thus for every non-leaf node x in T_n, chain C_x is the left boundary of a face in $P(T_n)$ that can be quadrangulated to add $|C_x| - 2$ extra faces to $P(T_n)$. We wish to count the number F_n of such faces added:

$$F_n = \sum_{\text{nonleaf nodes } x \text{ in } T_n} (|C_x| - 2). \tag{3}$$

Observe that if $x \in N_n(d)$, then $|C_x| = n - 2d$, since $|C_{0^n}| = n$ and for $y \neq 0^n$ in T_n, $|C_{p(y)}| = |C_w| + 2$. Then

$$F_n = \sum_{d=0}^{\lfloor \frac{n}{2} \rfloor} (n - 2d - 2)\overline{L}_n(d). \tag{4}$$

Our goal is to show that $F_n = 2^{n-1} - \binom{n}{\lfloor \frac{n}{2} \rfloor}$. We first need to be able to compute $\overline{L}_n(d)$.

Lemma 1 *If $n > 1$ and $0 \leq d \leq \lfloor \frac{n-1}{2} \rfloor$, then $\overline{L}_n(d) = \binom{n-1}{d} - \binom{n-1}{d-1}$, and 0 otherwise.*

Proof. We show that $\overline{L}_n(d) = N_{n-1}(d)$ by showing that the mapping $f(x) = x0$ is a bijection from the set of nodes in T_{n-1} of weight d to the set of non-leaves in T_n of weight d. The result follows then from (2).

If x is a node of weight d in T_{n-1}, x has no unmatched 1, so neither does $x0$, so $x0$ is a node of weight d in T_n. To see that $x0$ is a nonleaf, note that since $2d < n$, x has an unmatched 0, so $x1$ is in T_n and $p(x1) = x0$. Clearly f is one-to-one. To see that f is onto, let y be a nonleaf node of weight d in T_n. Then y has a child z and z has the form $z = \alpha 10^i$ for some $i \geq 0$. Since z is in T_n, α has no unmatched 1 and $y = p(\alpha 10^i) = \alpha 0^{i+1}$. Thus, $\alpha 0^i$ is in T_{n-1} and $y = f(\alpha 0^i)$. □

Lemma 2 *For $n \geq 2$,*

$$\sum_{d=0}^{\lfloor \frac{n}{2} \rfloor} \overline{L}_n(d) = \binom{n-1}{\lfloor \frac{n}{2} \rfloor - 1}.$$

Proof. Note that the summand is 0 if $d = \lfloor n/2 \rfloor$. Applying Lemma 1 gives a telescoping sum. □

Lemma 3 *For $n \geq 2$,*

$$\sum_{d=0}^{\lfloor \frac{n}{2} \rfloor} d\overline{L}_n(d) = \begin{cases} \frac{n}{2}\binom{n-1}{\frac{n}{2}-1} - 2^{n-2} & \text{if } n \text{ is even} \\ \frac{n-1}{2}\binom{n-1}{\lfloor \frac{n}{2} \rfloor - 1} + \frac{1}{2}\binom{n-1}{\lfloor \frac{n}{2} \rfloor} - 2^{n-2} & \text{if } n \text{ is odd.} \end{cases}$$

Proof. First note that the summand is 0 when $d = 0$ or $d = \lfloor \frac{n}{2} \rfloor$ and then apply Lemma 1 to get

$$\sum_{d=0}^{\lfloor \frac{n}{2} \rfloor} d\overline{L}_n(d) = \sum_{d=1}^{\lfloor \frac{n}{2} \rfloor - 1} d\left(\binom{n-1}{d} - \binom{n-1}{d-1}\right) \tag{5}$$

$$= \sum_{d=1}^{\lfloor \frac{n}{2} \rfloor - 1} \left(d\binom{n-1}{d} - (d-1)\binom{n-1}{d-1}\right) - \sum_{d=1}^{\lfloor \frac{n}{2} \rfloor - 1} \binom{n-1}{d-1}. \tag{6}$$

The first sum telescopes, giving $(\lfloor \frac{n}{2} \rfloor - 1)\binom{n-1}{\lfloor \frac{n}{2} \rfloor - 1}$. The second sum is the number of elements in the first $\lfloor \frac{n}{2} \rfloor - 2$ levels of the Boolean lattice $\mathcal{B}_{n-1}$ which, when n is even, is:

$$\frac{1}{2}\left(2^{n-1} - 2\binom{n-1}{\lfloor \frac{n}{2} \rfloor - 1}\right),$$

and when n is odd:

$$\frac{1}{2}\left(2^{n-1} - 2\binom{n-1}{\lfloor \frac{n}{2} \rfloor - 1} - \binom{n-1}{\lfloor \frac{n}{2} \rfloor}\right).$$

□

Theorem 4 *For all $n \geq 2$, the number of faces in the plane graph $P(T_n)$ after quadrangulating is 2^{n-1}. Thus, there are 2^{n-1} vertices in its dual, which is a Venn diagram for n sets.*

Proof. The embedding of the chain cover graph has $\binom{n}{\lfloor n/2 \rfloor}$ faces. We show that the number of such edges added in the quadrangulation phase is $F_n = 2^{n-1} - \binom{n}{\lfloor n/2 \rfloor}$. By (4),

$$F_n = \sum_{d=0}^{\lfloor \frac{n}{2} \rfloor} (n - 2d - 2)\overline{L}_n(d) = (n-2)\sum_{d=0}^{\lfloor \frac{n}{2} \rfloor} \overline{L}_n(d) - 2\sum_{d=0}^{\lfloor \frac{n}{2} \rfloor} d\overline{L}_n(d). \tag{7}$$

Using Lemmas 2 and 3, when n is even (7) becomes

$$(n-2)\binom{n-1}{\frac{n}{2}-1} + 2^{n-1} - n\binom{n-1}{\frac{n}{2}-1} = 2^{n-1} - 2\binom{n-1}{\frac{n}{2}-1} = 2^{n-1} - \binom{n}{\frac{n}{2}}.$$

When n is odd (7) becomes

$$\begin{aligned}(n-2)\binom{n-1}{\lfloor\frac{n}{2}\rfloor-1}+2^{n-1}-(n-1)\binom{n-1}{\lfloor\frac{n}{2}\rfloor-1}-\binom{n-1}{\lfloor\frac{n}{2}\rfloor} &= 2^{n-1}-\binom{n-1}{\lfloor\frac{n}{2}\rfloor-1}-\binom{n-1}{\lfloor\frac{n}{2}\rfloor} \\ &= 2^{n-1}-\binom{n}{\lfloor\frac{n}{2}\rfloor}.\end{aligned}$$

□

This number is one more than half the number of vertices in a simple diagram of order n. Thus we propose to call these diagrams “half-simple”.

5 At Least Half-Simple Symmetric Diagrams for Prime n

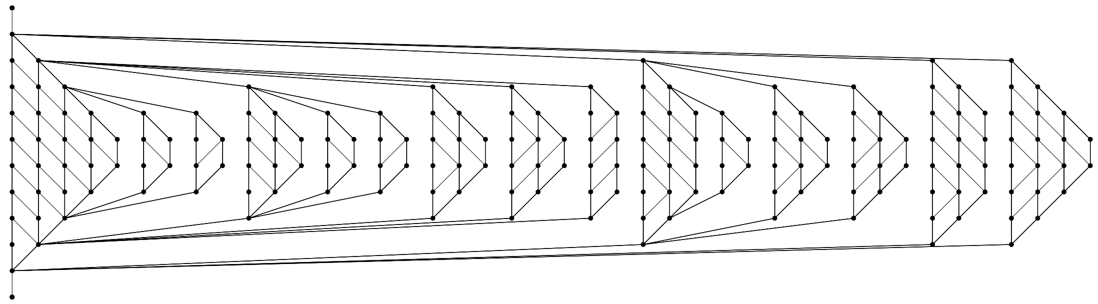
Recall that in Section 2, for prime n, the plane graph $R(P(S_n))$ was created from n copies of $P(S_n)$ and the dual of $R(P(S_n))$ was a symmetric Venn diagram with $\binom{n}{\lfloor n/2\rfloor}$ vertices. By Theorem 3, we can quadrangulate every face corresponding to a node and its first child in S_n *in every copy of* $P(S_n)$ and the resulting quadrangulation of $R(P(S_n))$ is still a plane, monotone, symmetric, spanning subgraph of the n-cube; thus its dual is still a symmetric Venn diagram. In this section we show that the total number of faces in the quadrangulation of $R(P(S_n))$ is at least $2^{n-1}(1-o(1))$ and therefore its dual is asymptotically at least half-simple.

As an example for $n = 11$ (see Figure 10), the number of faces of $P(S_{11})$ before quadrangulating is $\binom{11}{5}/11 = 42$. The number of faces added by quadrangulating the faces between the chains corresponding to a node in S_{11} and its first child is 69. Repeating this in every copy of $P(S_{11})$, the total number of faces in the resulting quadrangulation of $R(P(S_{11}))$ is 11(42+69)=1221. In addition, note that any circular permutation can be used to label subsequent rotated copies of $P(S_{11})$; if the circular permutation used is (6 7 8 9 10 11 1 2 3 4 5), then 16 extra faces can be added to $P(S_{11})$ by manually adding extra edges wherever possible, giving another $11 \times 16 = 171$ faces in the graph. Thus, the dual is a symmetric Venn diagram with $1221+171 = 1392$ vertices, whereas a simple Venn diagram would have 2046 vertices. This diagram is rigid, *i.e.* as simple as possible, as no more edges can be added to the dual graph and thus no more vertices can be separated in the Venn diagram.

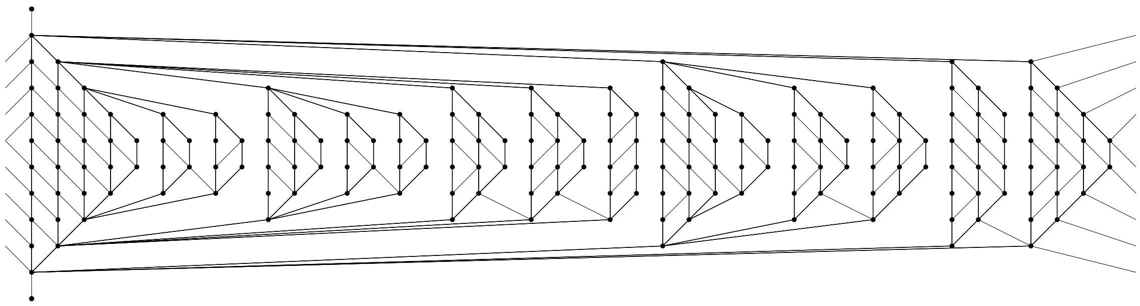
For contrast, in [6], Hamburger shows how to separate vertices in his 11-Venn diagram from [7] to get symmetric 11-Venn diagrams with only up to 1001 vertices. Following his example, we can say that since each of the $69 + 16 = 85$ extra edges can either be present or not in the rotated copies of $P(S_{11})$, these extra edges give us 2^{85} distinct symmetric 11-Venn diagrams, more than previously known.

It remains to count F_n, the number of faces added to $P(S_n)$ by quadrangulating the faces between the chains corresponding to a node in S_n and its first child. Then nF_n is the total number of faces added to $R(P(S_n))$.

Recall from Section 2 that x is a node in S_n if and only if



(a) After quadrangulating all faces corresponding to a node in S_{11} and its first child.



(b) After manually adding extra edges, including wrapping edges.

Figure 10: Adding edges to the graph $P(S_{11})$.

- (i) x has finite block code (so it starts with 1 and ends with 0),
- (ii) $\beta(x) \leq \beta(\sigma^i(x))$ for any i, and
- (iii) x has exactly one unmatched 1.

We will make use of the following lemma:

Lemma 4 *If x is a node in the chain cover tree and if the last block of x has the form 10^k, then $k \geq 2$.*

Proof. First, $k \geq 1$ as $\beta(x)$ is finite. If $k = 1$, then $\beta(x) = (\ldots, 2)$. However as $n = |x|$ is prime there must be an element in $\beta(x)$, call it j, such that $j \geq 3$ and so we could create a rotation of x with a lexicographically smaller block code by rotating two positions right to create x' with $\beta(x') = (2, \ldots)$. This contradicts the fact that x has the (unique) lexicographically lowest block code of all of its rotations. □

Let $\ell(x)$ denote the location of the last 1 in x.

Theorem 5 *If u and v are children of w in the chain cover tree and $\ell(u) < \ell(v)$, then v is not a leaf.*

Proof. Since w satisfies (i), w can be written as

$$w = \alpha 1^a 0^b,$$

where either α is empty or α has finite block code. Then u and v have the form

$$u = \alpha 1^a 0^c 1 0^{b-c-1}; \qquad v = \alpha 1^a 0^{c'} 1 0^{b-c'-1},$$

where $0 \le c = \ell(u) - |\alpha| - a - 1$ and $c < c' = \ell(v) - |\alpha| - a - 1$. Let

$$z = \alpha 1^a 0^{c'} 1 1 0^{b-c'-2}.$$

If we show that z satisfies (i)-(iii), then z is a child of v in the chain cover tree.

Applying Lemma 4 to v gives that $b - c' - 2 \ge 2$, so z ends in 0 and starts with 1, satisfying (i). Also $\beta(z) = \beta(v)$, so z satisfies (ii). Finally, since u is a node, it satisfies (iii) and therefore the following string has exactly one unmatched 1, which, since $c \ge 1$, is not the last 1:

$$\alpha 1^a 0^c 1.$$

Then, since $c' > c$, the string

$$\alpha 1^a 0^{c'} 1$$

has exactly one unmatched 1 and at least one unmatched 0, so the string

$$\alpha 1^a 0^{c'} 1 1$$

(and therefore also z) has exactly one unmatched 1. So z satisfies (iii). $\square$

Corollary 1 *Each node in the chain cover tree has at most one leaf child.*

We make the following definitions.

- Let $N(d)$ denote the number of nodes of weight d in S_n. Then $N(1) = 1$ and for $d > 1$,
$$N(d) = \frac{1}{n}\left[\binom{n}{d} - \binom{n}{d-1}\right], \tag{8}$$
- Let $L(d)$ denote the number of leaves of weight d in S_n.
- Let $\overline{L}(d)$ denote the number of nonleaves of weight d in S_n.
- Let $w(d)$ denote the number of faces added to $P(S_n)$ by quadrangulating the region between the chains corresponding to a node x of weight d its first child w in S_n. ¿From Theorem 2 we know that
$$w(d) = |C_w| = |C_x| - 2 = n - 2d - 2 \quad \text{for } 1 \le d \le (n-3)/2. \tag{9}$$

Then F_n, the number of faces added to $P(S_n)$, can be expressed as

$$F_n = \sum_{d=1}^{(n-3)/2} w(d)\overline{L}(d).$$

We can get a lower bound on F_n by observing first that every node at level d contributes either $w(d-1)$ toward quadrangulating a region with its parent (if it is a leaf) or $w(d)$ toward quadrangulating a region with its first child (if it is a nonleaf) or both. By Corollary 1, leaves can be mapped one-to-one to parents. Thus if we count all of the contributions at every node at every level, we are at worst double-counting, giving the second inequality below. The first inequality follows since $w(d) \le w(d-1)$ and $N(d) = L(d) + \overline{L}(d)$.

$$\sum_{d=1}^{(n-3)/2} w(d)N(d) \le \sum_{d=1}^{(n-3)/2} (w(d-1)L(d) + w(d)\overline{L}(d) \le 2 \sum_{d=1}^{(n-3)/2} \overline{L}(d)w(d).$$

Thus, the number of faces added to $R(P(S_n))$ by quadrangulating is

$$nF_n \ge \frac{n}{2} \sum_{d=1}^{(n-3)/2} w(d)N(d). \tag{10}$$

Theorem 6 *For n prime, the number of faces added to $R(P(S_n))$ by quadrangulating is*

$$nF_n \ge 2^{n-1} - 1 - \frac{1}{2}\left[\binom{n+1}{(n+1)/2} + \binom{n}{(n-3)/2}\right], \tag{11}$$

making its dual a symmetric Venn diagram with $2^{n-1}(1-o(1))$ vertices.

Proof. We need to show that (11) is a lower bound for (10).

We will make use of the following identity, established in the proof of Theorem 4, with the observation that in (7), $\overline{L}(\lfloor \frac{n}{2} \rfloor) = 0$.

$$\sum_{d=0}^{\lfloor \frac{n}{2} \rfloor - 1} (n-2d-2)\left[\binom{n-1}{d} - \binom{n-1}{d-1}\right] = 2^{n-1} - \binom{n}{\lfloor n/2 \rfloor}. \tag{12}$$

If we replace n by $n+1$ and observe that the summand on the left is $n-1$ when $d=0$ then for odd n, we get

$$\sum_{d=1}^{(n+1)/2-1} (n-2d-1)\left[\binom{n}{d} - \binom{n}{d-1}\right] = 2^n - \binom{n+1}{(n+1)/2} - (n-1). \tag{13}$$

Starting from (10) and using (8) and (9),

$$\begin{aligned}
\frac{n}{2}\sum_{d=1}^{(n-3)/2} w(d)N(d) &= \frac{n}{2}\left((n-4)+\sum_{d=2}^{(n-3)/2}(n-2d-2)\frac{1}{n}\left[\binom{n}{d}-\binom{n}{d-1}\right]\right)\\
&= \frac{1}{2}\left(n(n-4)-(n-1)(n-4)+\sum_{d=1}^{(n-3)/2}(n-2d-2)\left[\binom{n}{d}-\binom{n}{d-1}\right]\right)\\
&= \frac{1}{2}\left(n-4+\sum_{d=1}^{(n-3)/2}(n-2d-1)\left[\binom{n}{d}-\binom{n}{d-1}\right]\right.\\
&\qquad\left.-\sum_{d=1}^{(n-3)/2}\left[\binom{n}{d}-\binom{n}{d-1}\right]\right)\\
&= \frac{1}{2}\left(n-4+\sum_{d=1}^{(n-3)/2}(n-2d-1)\left[\binom{n}{d}-\binom{n}{d-1}\right]-\binom{n}{(n-3)/2}+1\right).
\end{aligned}$$

We now make use of (13) to evaluate the sum in the last line above, noting that the last term in the sum on the left of (13) is 0:

$$\begin{aligned}
\frac{n}{2}\sum_{d=1}^{(n-3)/2} w(d)N(d) &= \frac{1}{2}\left(n-4+2^n-\binom{n+1}{(n+1)/2}-(n-1)-\binom{n}{(n-3)/2}+1\right)\\
&= 2^{n-1}-1-\frac{1}{2}\left(\binom{n+1}{(n+1)/2}+\binom{n}{(n-3)/2}\right).
\end{aligned}$$

This lower bound asymptotically approaches 2^{n-1} as n increases. □

In fact, the number of faces resulting from the quadrangulations appears to be considerably larger than the lower bound given by the theorem, but we have been unable to find a tighter counting argument which would establish this. The table below shows a comparison of the number of vertices in a simple Venn diagram, the number of vertices in the symmetric Venn diagrams produced by the GKS construction, the number of vertices in the symmetric Venn diagrams produced by the construction of this paper (KRSW), and the ratio of the number of vertices produced by the KRSW construction to the number in a simple Venn diagram.

n	simple	[GKS]	[KRSW]	$\frac{\text{[KRSW]}}{\text{simple}}$
3	6	3	3	.5000
5	30	10	15	.5000
7	126	35	70	.5556
11	2,046	462	1,221	.5968
13	8,190	1,716	5,005	.6111
17	131,070	24,310	81,787	.6240
19	524,286	92,378	329,289	.6281
23	8,388,606	1,352,078	5,308,423	.6328

6 Concluding Remarks

We have succeeding in showing that for prime n there are symmetric Venn diagrams that are nearly simple, in the sense that the average number of curves passing through a point of intersection is at most a constant c, independent of n. Although we have established this for (asymptotically) $c = 4$, the same construction with an improved counting argument could likely lower the constant c. The question remains as to whether there are simple symmetric Venn diagrams for n prime and greater than 7; even the the $n = 11$ case remains open.

It would also be interesting to prove an upper bound on the number of symmetric Venn diagrams.

References

[1] Bette Bultena, Branko Grünbaum, and Frank Ruskey. Convex drawings of intersecting families of simple closed curves. In *11th Canadian Conference on Computational Geometry*, pages 18–21, 1999.

[2] Bette Bultena and Frank Ruskey. Venn diagrams with few vertices. *Electron. J. Combin.*, 5(1):Research Paper 44, 21 pp. (electronic), 1998.

[3] Curtis Greene and Daniel J. Kleitman. Strong versions of Sperner's theorem. *J. Combinatorial Theory Ser. A*, 20(1):80–88, 1976.

[4] Jerrold Griggs, Charles E. Killian, and Carla D. Savage. Venn diagrams and symmetric chain decompositions in the Boolean lattice. *Electron. J. Combin.*, 11:Research Paper 2, 30 pp. (electronic), 2004.

[5] Branko Grünbaum. Venn diagrams and independent families of sets. *Math. Mag.*, 48:12–23, 1975.

[6] Peter Hamburger. Pretty drawings. More doodles and doilies, symmetric Venn diagrams. *Utilitas Mathematica*. To appear, preprint (http://www.ipfw.edu/math/hamburger/papers.htm).

[7] Peter Hamburger. Doodles and doilies, non-simple symmetric Venn diagrams. *Discrete Math.*, 257(2-3):423–439, 2002.

[8] D. W. Henderson. Venn diagrams for more than four classes. *American Mathematical Monthly*, 70:424–426, 1963.

[9] Frank Ruskey. A survey of Venn diagrams. *Electron. J. Combin.*, 4(1):Dynamic Survey 5 (electronic), 2001.